Collezioni digitali di periodici in Italia e in Europa: standard, applicazioni, valutazioni, prospettive

Arco (TN) - *15-16 Novembre* 2007

Come contare gli utenti? Le basi tecnologiche per rilevare gli utenti di un sito

Zeno Tajoli – tajoli@cilea.it

Abstract

In questo workshop si vogliono illustrare le basi e le problematiche connesse alla rilevazione degli utenti di un sito, le loro caratteristiche, le operazioni che compiono.
Per compiere queste rilevazioni si illustrera' come funziona in generale la comunicazione tra un utente e un sito e come gli utenti siano anonimi, ma lascino sempre dei dati, che una volta interpretati, possono dire molto di loro.
In particolare verranno illustrate le problematiche sottese a queste operazioni,per permettere al bibliotecario di capire meglio cosa significano esattamente i dati forniti dai software che valutano l'uso di risorse digitali accessibili a tutti gli utenti di Internet.

Prerequisiti

- Uso di Internet
- Generica conoscenza dei PC
- Interesse a guardare dietro le quinte
- Nessun vero background tecnico richiesto
- Si parte dalla base

Punti salienti

- Come si comunica?
- Che informazioni ho?
- Come estrarre i dati.
- Superare i limiti presentati
- Un esempio

Come si comunica ?

Sono necessari meccanismi software per permettere ai vari computer di dialogare di gestire la comunicazione

- protocolli (convenzioni) di comunicazione
- meccanismi di indirizzamento (come identificare un computer)
- spedizione sulle connessioni opportune

Come si comunica ?

- invio e ricezione di messaggi
- verifica correttezza dei messaggi durante la trasmissione
- protezione dei messaggi (per evitare intercettazione)
- ottimizzazione della comunicazione
- gestione del traffico sulla rete

Come si comunica ?

Un protocollo umano e un protocollo di reti di computer:
1) Ciao
2) Ciao
3) Hai l'ora?
4) 2:00

a) TCP connection
b) request
c) TCP connection
d) reply.
e) Get http://www.di.unito.it/index.htm
f) <file>

Domanda: Altri protocolli umani?

Come si comunica ?

Protocolli umani:

- "Che ora è?"
- "Ho una domanda"
- Presentazioni…

… messaggi specifici vengono spediti

… azioni specifiche sono compiute quando i messaggi sono ricevuti, o in seguito ad altri eventi

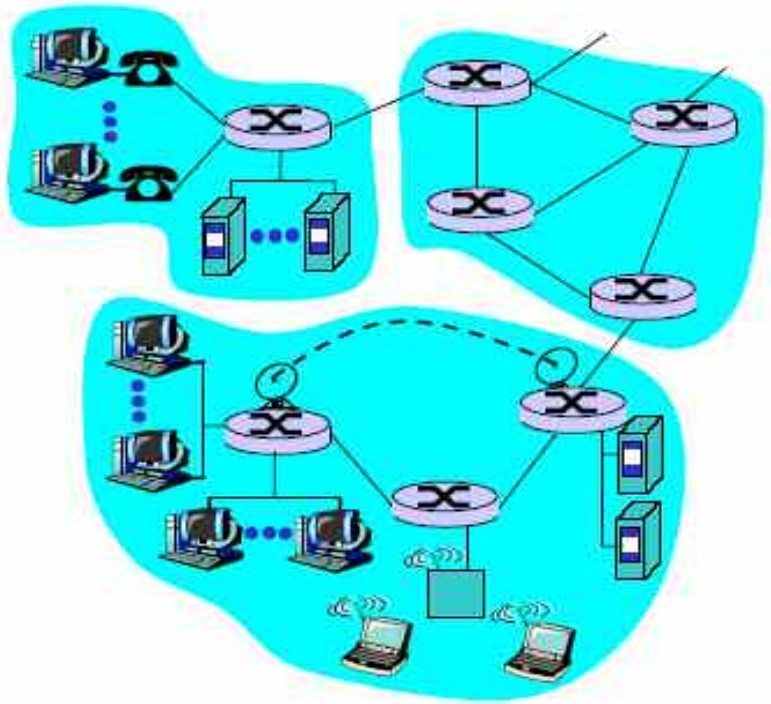
Protocolli di rete:

- macchine invece di esseri umani
- Tutte le attività di comunicazione in Internet sono governate da protocolli

I protocolli definiscono formato e ordine dei messaggi spediti e ricevuti tra entità della rete, e le azioni da compiere in seguito alla ricezione e/o trasmissione dei messaggi o di altri eventi

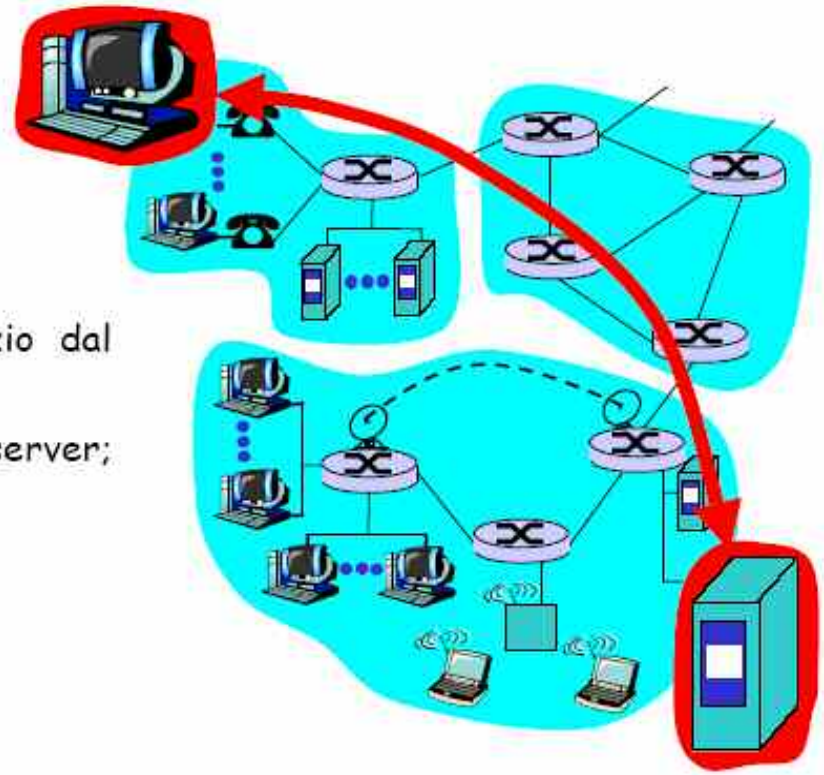
Come si comunica ?

- network edge: applicazioni ed host
- network core:
  - router
  - rete di reti
- reti di accesso, mezzi trasmissivi: canali di comunicazione



Come si comunica ?

- end systems (host):
  - Eseguono programmi applicativi
  - e.g., WWW, email
  - al "bordo della rete"
- modello client/server
  - il client richiede, riceve servizio dal server
  - e.g., WWW client (browser)/ server; email client/server
- modello peer-peer:
  - interazione tra host simmetrica
  - e.g.: Gnutella, KaZaA



Come si comunica ?

Obiettivo: trasferimento dati tra host

- *handshaking:* fase di preparazione antecedente al trasferimento dati
  - Ciao - Ciao nel protocollo umano
  - *Stabilire uno "stato"* nei due host comunicanti
- TCP - Transmission Control Protocol
  - Servizio di scambio dati di tipo connection-oriented di Internet

Servizio TCP [RFC 793]

- *Trasferimento affidabile ed ordinato di byte di un flusso dati*
  - perdite: conferma di ricezione (acknowledgement) e ri-trasmissione
- *Controllo di flusso*
  - Il mittente non sovraccaricherà il ricevitore
- *Controllo di congestione:*
  - I mittenti diminuiscono la loro velocità di spedizione quando la rete si congestiona

Come si comunica ?

Obiettivo: trasferimento dati tra host

- Esattamente lo stesso!

- UDP - User Datagram Protocol [RFC 768]: Servizio connectionless di Internet
  - Senza handshaking
  - Trasferimento dati non-affidabile
  - senza controllo di flusso
  - senza controllo congestione

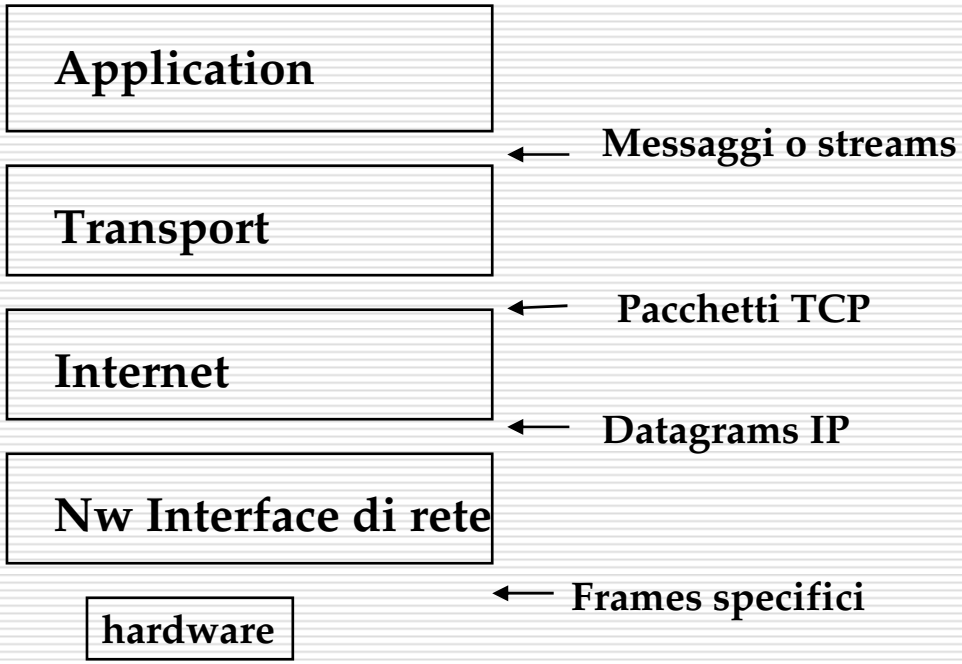
Applicazioni che usano TCP:

- HTTP (WWW), FTP (trasferimento file), Telnet (login remoto), SMTP (email)

Applicazioni che usano UDP:

- streaming media, teleconferencing, Internet telephony

Come si comunica ?



Il www in realtà usa due protocolli, il TCP e l'IP

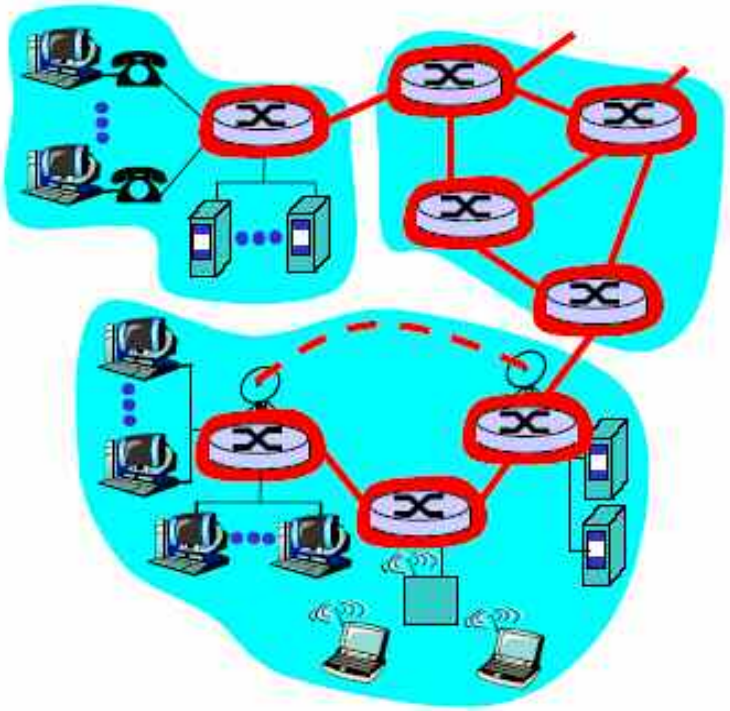
Come si comunica ?

Il protocollo IP

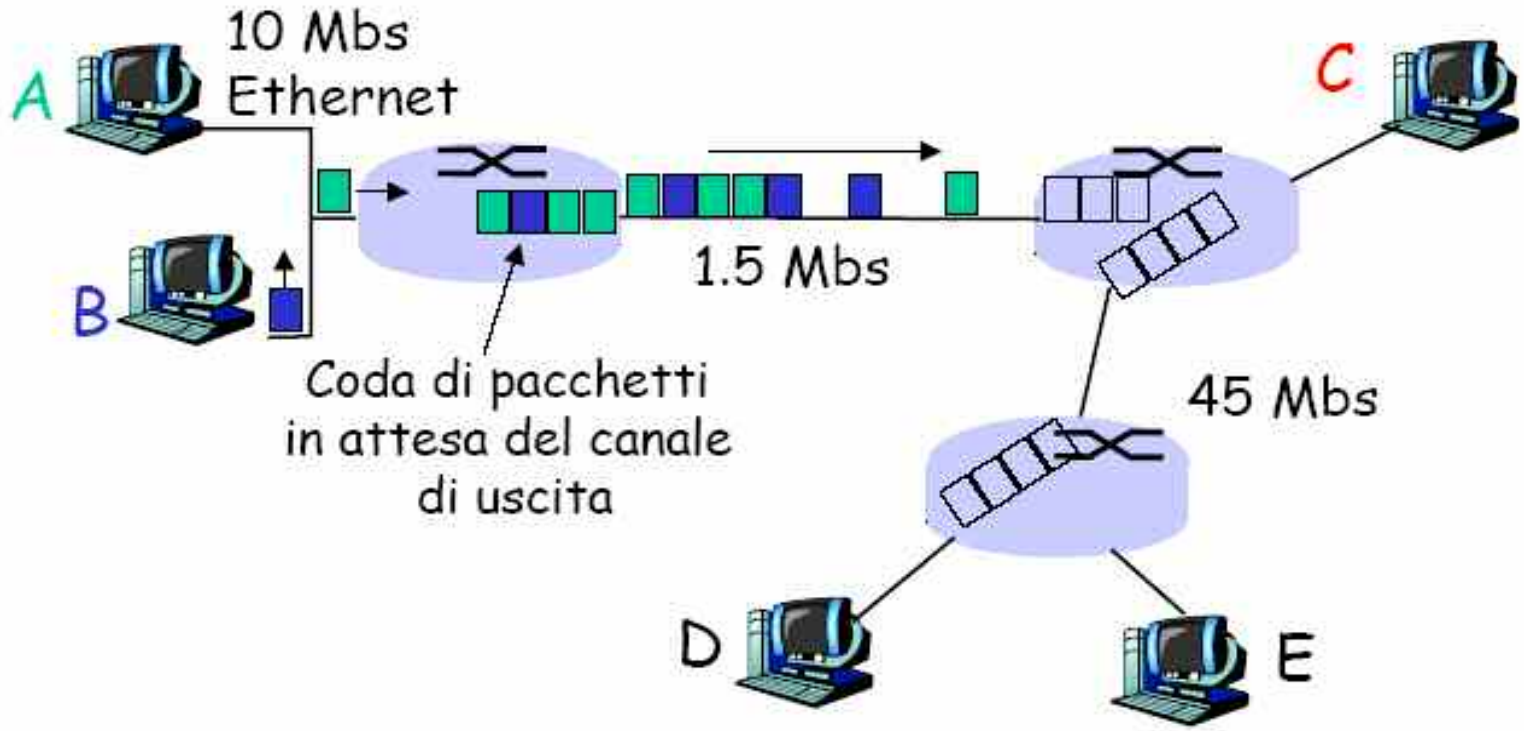
- protocollo che lavora con pacchetti (datagram) indipendenti e di formato predefinito
- i nodi possono eliminare datagram se mancano le risorse
- nessun algoritmo di correzione sui datagram
- frammentazione lungo la strada e riassemblaggio gestito alla destinazione
- nessuna garanzia di instradamento dei datagram
- riconfigurazione automatica della strada percorsa in caso di malfunzionamenti
- distruzione dei datagram se ripetitivi

Come si comunica ?

- Maglia di router interconnessi
- *Domanda fondamentale*: come vengono trasferiti i dati attraverso la rete?
  - Commutazione di pacchetto: i dati sono spediti attraverso la rete in quantità discrete chiamate pacchetti



Come si comunica ?



Come si comunica ?

Ogni flusso dati viene diviso in *pacchetti*

- I pacchetti degli utenti A e B *condividono* risorse di rete
- Ogni pacchetto usa tutta la larghezza di banda (capacità di trasmissione in bit al secondo) del canale
- Risorse usate quando sono necessarie

Contesa delle risorse:

- La richiesta aggregata di risorse può eccedere l'ammontare disponibile
- congestione: i pacchetti si accodano ed attendono l'uso del canale
- store and forward: pacchetti ricevuti interamente prima di essere spediti

Come si comunica ?

- *Obiettivo:* spostare pacchetti tra router, dal host sorgente all' host destinatario
- Caratteristiche:
  - *L'indirizzo destinazione* determina il prossimo passo
  - Le strade (route) possono variare durante le sessioni
  - I router NON mantengono informazioni sullo stato delle connessioni

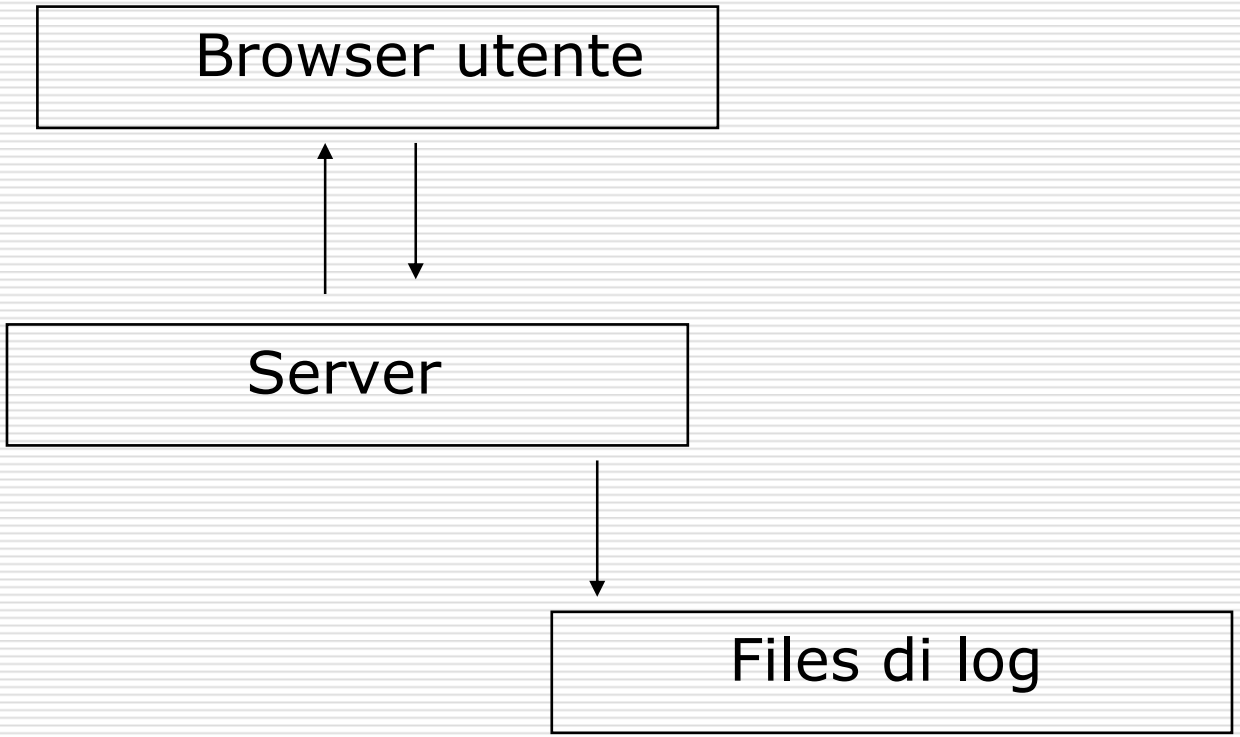
Come si comunica ?

- Una struttura di comunicazione complessa
- Dimensionabile da 2 a n sistemi
- Senza preconoscenza di tutta la struttura
- Senza totale anonimato
- Senza totale sicurezza di chi è e cosa ha realmente ricevuto

Come si comunica ?

- La comunicazione è tra macchine
- Il dato fondamentale è l' IP
- Ma non tutti l'hanno fisso
- I grandi fornitori di connettività (TIM, AOL, Tiscali, etc.) ne hanno un pool che fanno usare, girandoli, tra gli utenti finali
- Anche server con IP dinamico (Dynamic DNS)

Che informazioni ho ?



Che informazioni ho ?

- Via via che la comunicazione client/server va avanti, il server scrive quanto avviene su un file
- Di norma i file sono due,
  - Il primo per le operazioni normali
  - Il secondo per quanto ha causato errori
- Si analizzerà quanto fa il server web Apache 2, simile é il comportamento di IIS di Microsoft

Che informazioni ho ?

- 213.58.139.xx "-" "-" [11/Nov/2007:06:43:13 +0100] "GET /archive/00011300/ HTTP/1.1" 200 12384 "http://www.google.pt/search?q=caf&hl=pt-PT&lr=lang_pt&start=170&sa=N" "Mozilla/4.0(MSIE 6.0)"
- 213.58.139.xx "-" "-" [11/Nov/2007:06:43:13 +0100] "GET /eprints.css HTTP/1.1" 200 3723 "http://eprints.rclis.org/archive/00011300/" " Mozilla/4.0(MSIE 6.0)" "
- 213.58.139.xx "-" "-" [11/Nov/2007:06:43:13 +0100] "GET /images/imatge.jpg HTTP/1.1" 200 38180 "http://eprints.rclis.org/archive/00011300/" " Mozilla/4.0(MSIE 6.0)"
- 213.58.139.xx "-" "-" [11/Nov/2007:06:43:13 +0100] "GET /images/mon.jpg HTTP/1.1" 200 14692 "http://eprints.rclis.org/archive/00011300/" "Mozilla/4.0(MSIE 6.0)"
- 213.58.139.xx "-" "-" [11/Nov/2007:06:43:13 +0100] "GET /images/dibuix.jpg HTTP/1.1" 200 781 "http://eprints.rclis.org/archive/00011300/" "Mozilla/4.0(MSIE 6.0)"
- 213.58.139.xx "-" "-" [11/Nov/2007:06:43:13 +0100] "GET /images/reflink.png HTTP/1.1" 200 353 "http://eprints.rclis.org/archive/00011300/" "Mozilla/4.0(MSIE 6.0)"
- 213.58.139.xx "-" "-" [11/Nov/2007:06:43:13 +0100] "GET /email.js HTTP/1.1" 200 194 "http://eprints.rclis.org/archive/00011300/" " Mozilla/4.0(MSIE 6.0)"
- 72.226.75.yy "-" "-" [11/Nov/2007:06:43:14 +0100] "GET /7876/ HTTP/1.1" 200 5 "-" "Mozilla/5.0 Firefox/2.0.0.9"

Che informazioni ho ?

- IP-address remoto
- IP-address locale
- Dimensione della risposta (senza gli header)
- Dati del cookie inviato dal browser
- Tempo per gestire la richiesta sul server
- Variabili interne del server
- Nome del file inviato
- Protocollo di richiesta
- Informazioni che il browser dà di se stesso
  - Che lingua richiede
  - Che set di caratteri usa
  - Che pagina ha visto prima
  - Come si chiama il software del browser
  - Altre informazioni molto variabili tra i diversi software
- LogNome dell'utente remoto se autenticato
- Il modo di fare la richiesta

Che informazioni ho ?

- Le intestazioni inviate nella risposta
- La porta di comunicazione
- I numeri interni del servizio nella memoria (PID)
- La parte domanda di una richiesta ricevuta
- La richiesta ricevuta
- Lo status della comunicazione
- L'ora in cui si è ricevuta la richiesta
- Nome associato al LogName dell'utente remoto se autenticato
- La richiesta ricevuta senza la eventuale parte domanda
- Nome del server
- Status della connessione dopo l'invio di quanto chiesto
- Bytes ricevuti tutto compreso
- Bytes inviati tutto compreso

Riferimento completi: RFC 2616 e modulo Apache 2 mod_log_config

Che informazioni ho ?

- 213.58.139.xx "-" "-" [11/Nov/2007:06:43:13 +0100] "GET /archive/00011300/ HTTP/1.1" 200 12384 "http://www.google.pt/search?q=caf&hl=pt-PT&lr=lang_pt&start=170&sa=N" "Mozilla/4.0(MSIE 6.0)"

Da IP 213.58.139.xx un browser MSIE chiama eprints.rclis.org e chiede la home di 11300, viene da una query su google, la richiesta è andata buon fine e sono stati inviati 12384 bytes

- 213.58.139.xx "-" "-" [11/Nov/2007:06:43:13 +0100] "GET /eprints.css HTTP/1.1" 200 3723 "http://eprints.rclis.org/archive/00011300/" " Mozilla/4.0(MSIE 6.0)"

Sempre allo stesso inviato con sucesso il file eprints.css

- 213.58.139.xx "-" "-" [11/Nov/2007:06:43:13 +0100] "GET /images/imatge.jpg HTTP/1.1" 200 38180 "http://eprints.rclis.org/archive/00011300/" " Mozilla/4.0(MSIE 6.0)"

Sempre allo stesso inviato con sucesso il file imatge.jpg

- 213.58.139.xx "-" "-" [11/Nov/2007:06:43:13 +0100] "GET /images/mon.jpg HTTP/1.1" 200 14692 "http://eprints.rclis.org/archive/00011300/" "Mozilla/4.0(MSIE 6.0)"

Sempre allo stesso inviato con sucesso il file mon.jpg

- Vari invii di diversi files allo stesso IP

Che informazioni ho ?

- 72.226.75.yy "-" "-" [11/Nov/2007:06:43:14 +0100] "GET /7876/ HTTP/1.1" 200 5 "-" "Mozilla/5.0 Firefox/2.0.0.9“

Connessione con sucesso di un altro IP

- I log di esempio hanno questa sintassi:
- "%h \"%l\" \"%u\" %t \"%r\" %>s %b \"%{Referer}i\" \"%{User-Agent}i\“”
- I dati sono in stretto ordine di tempo
- Dunque si accavallano l’un con l’altro

Come estraggo i dati.

- Ma quando finisce la sessione ?
- Posso collegare i vari passi dell'utente ?
- Se ritorna lo stesso IP è la stessa macchina ?
- Stessa macchina = stessa persona ?

Come estraggo i dati.

- I raggruppamenti e le successive analisi si fanno sfruttando di base:
  - Il valore dell'IP
  - L'orario dell'operazione
- La fine sessione viene definita attraverso un intervallo temporale minimo tra una nuova evenienza dello stesso IP
- Un limite dice che è lo stesso utente che fa una nuova sessione
- Un secondo limite dice che è un utente diverso.
- Un esempio:
  - Tra 0 e 5 minuti: stesso utente e stessa sessione
  - Tra 5 minuti e 1 ora: stesso utente ma diversa sessione
  - Otre 1 ora: diverso utente

Come estraggo i dati.

- La gestione dei limiti temporali è il passaggio cruciale le elaborazione.
- I dati sui files e su quanto è stato inviato sono dunque precisi e certi
- I dati sugli utenti sono dunque delle stime.
- Diversi software usano diversi algoritmi per queste stime

Come estraggo i dati.

- Stessa macchina non è stessa persona
- Si vogliono vedere le visite di una persona
- Dunque quando si supera il limite massimo, uno stesso utente o un utente diverso per noi pari sono.

Come estraggo i dati.

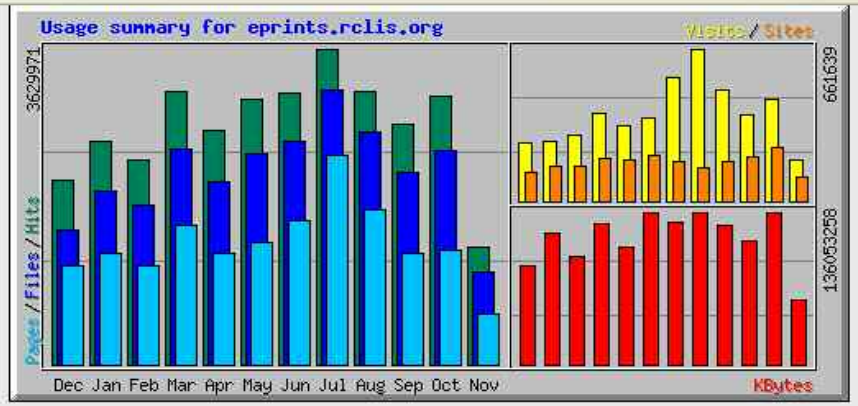
- Non sempre dietro una macchina c'è una persona
- Ci sono software che scaricano dal web per vari motivi (robots/spiders/ agents/etc.)
- Bisogna inferire dal loro comportamento che sono software
- Una volta identificati si riconoscono successivamente usando IP e nome del software di browsing
- La lista dei robots va costantemente aggiornata
- Vi sono anche scaricatori massivi usati dalle persone, i cosiddetti "site downloader"
- Di norma sono considerati come i robots

Usage Statistics for Eprints eprints.rclis.org - Last 12 Months - Mozilla Firefox

File Modifica Visualizza Cronologia Segnalibri Strumenti ?

http://eprints.rclis.org/statistics/index.html

Cilea: Home

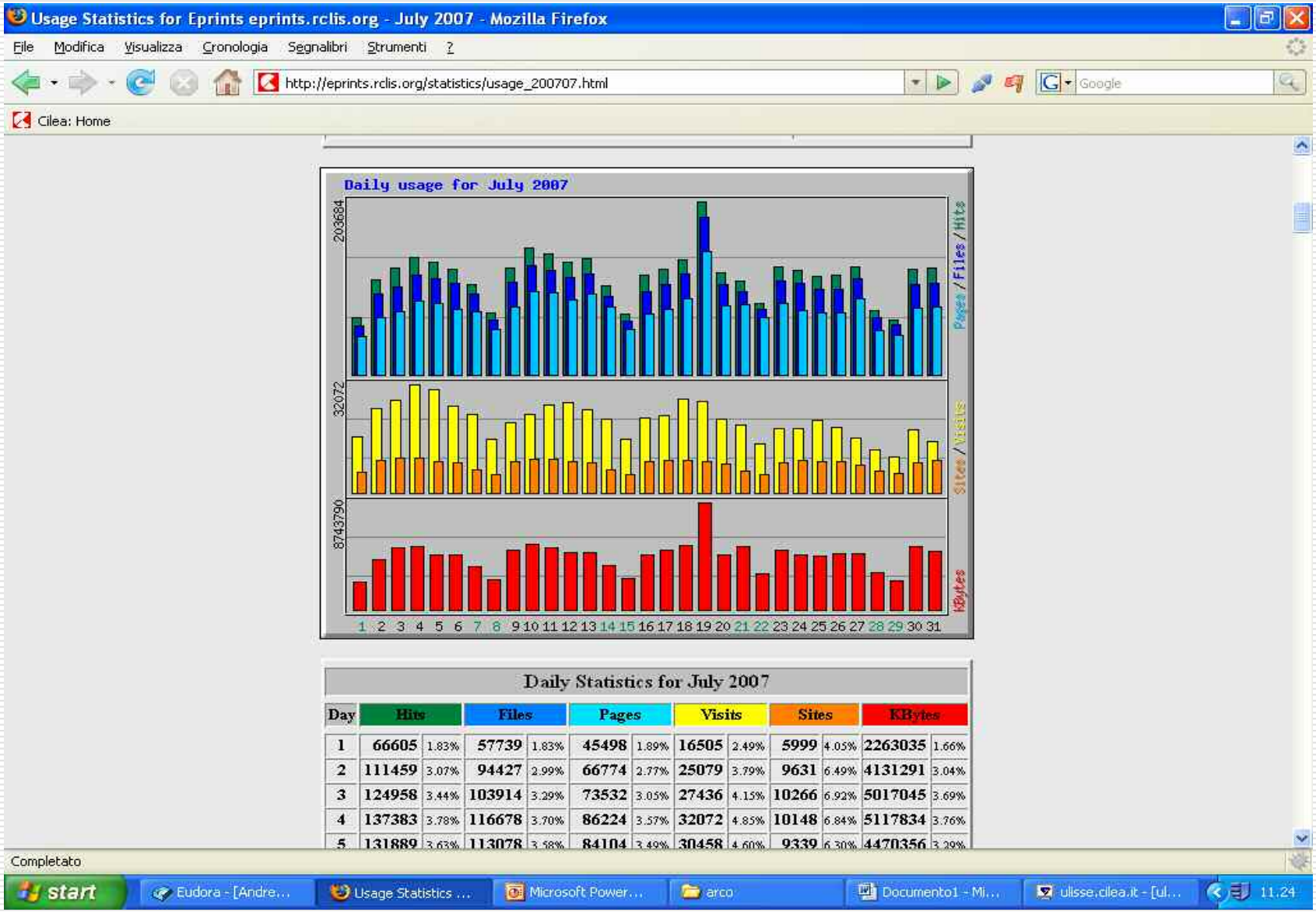


Summary by Month

Month	Daily Avg				Monthly Totals					
	Hits	Files	Pages	Visits	Sites	KBytes	Visits	Pages	Files	Hits
Nov 2007	103630	81249	44939	13759	105226	57284068	178868	584207	1056248	1347191
Oct 2007	99484	79559	42539	14345	231801	135102676	444719	1318718	2466351	3084026
Sep 2007	91798	73822	42377	12528	193708	109915983	375855	1271324	2214689	2753969
Aug 2007	101039	86059	57556	15570	172375	124310743	482673	1784263	2667844	3132216
Jul 2007	117095	101830	77810	21343	148301	136053258	661639	2412111	3156744	3629971
Jun 2007	103898	85547	55371	17923	172626	126985847	537704	1661140	2566412	3116943
May 2007	98530	78241	45004	11610	200805	134762624	359921	1395134	2425496	3054434
Apr 2007	89723	69853	42564	10834	176963	105394606	325044	1276944	2095617	2691699
Mar 2007	101149	79718	51885	12247	186454	125328858	379684	1608436	2471280	3135629
Feb 2007	83733	65757	40423	10323	152783	95985183	289061	1131849	1841210	2344524
Jan 2007	82964	64479	41253	8424	148631	117811573	261158	1278869	1998856	2571904
Dec 2006	68117	50092	36894	8145	123942	88421117	252506	1143726	1552863	2111640

Completato

start | Eudora - [Andre... | Usage Statistics ... | Microsoft Power... | arco | Documento1 - Mi... | ulisse.cilea.it - [ul... | 11.25

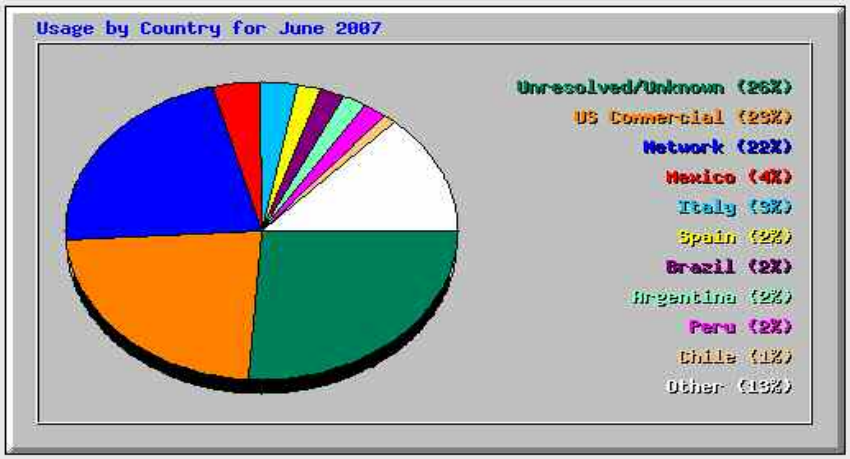

Usage Statistics for Eprints eprints.rclis.org - July 2007 - Mozilla Firefox
File Modifica Visualizza Cronologia Segnalibri Strumenti ?
http://eprints.rclis.org/statistics/usage_200707.html
Cilea: Home
Daily usage for July 2007
203684
32072
8743790
Pages / Files / Hits
Sites / Visits
KBytes
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31
Daily Statistics for July 2007
Day Hits Files Pages Visits Sites KBytes
1 66605 1.83% 57739 1.83% 45498 1.89% 16505 2.49% 5999 4.05% 2263035 1.66%
2 111459 3.07% 94427 2.99% 66774 2.77% 25079 3.79% 9631 6.49% 4131291 3.04%
3 124958 3.44% 103914 3.29% 73532 3.05% 27436 4.15% 10266 6.92% 5017045 3.69%
4 137383 3.78% 116678 3.70% 86224 3.57% 32072 4.85% 10148 6.84% 5117834 3.76%
5 131889 3.63% 113078 3.58% 84104 3.49% 30458 4.60% 9339 6.30% 4470356 3.29%
Completato
start
Eudora - [Andre...
Usage Statistics ...
Microsoft Power...
arco
Documento1 - Mi...
ulisse.cilea.it - [ul...
11.24

Usage Statistics for Eprints eprints.rclis.org - June 2007 - Mozilla Firefox

File Modifica Visualizza Cronologia Segnalibri Strumenti ?

http://eprints.rclis.org/statistics/usage_200706.html

Cilea: Home



Top 30 of 153 Total Countries

#	Hits		Files		KBytes		Country
1	805068	25.83%	736046	28.68%	28282525	22.27%	Unresolved/Unknown
2	727675	23.35%	661152	25.76%	25225476	19.86%	US Commercial
3	671027	21.53%	608976	23.73%	30348626	23.90%	Network
4	125165	4.02%	115282	4.49%	7000329	5.51%	Mexico
5	94312	3.03%	88096	3.43%	3980394	3.13%	Italy
6	74952	2.40%	60834	2.37%	3999428	3.15%	Spain
7	70067	2.25%	64367	2.51%	3410972	2.69%	Brazil
8	54168	1.74%	47082	1.83%	2602599	2.05%	Argentina
9	46912	1.51%	43588	1.70%	1907063	1.50%	Peru
10	32051	1.03%	27680	1.08%	2263872	1.78%	Chile

Completato

start | Eudora - [Andre... | Usage Statistics ... | Microsoft Power... | arco | Documento1 - Mi... | ulisse.cilea.it - [ul... | 11.25

Come estraggo i dati.

Diversi soft. su stessi dati = diverse stime

- **WEBALIZER**: http://www.webalizer.org
- **ANALOG**: http://www.analog.cx/
- **AWSTATS**: http://awstats.sourceforge.net/

Sono software generali che vanno bene per ogni sito

Come estraggo i dati.

Localizzare tramite l'IP

- Possibile tramite studi su quanto riferito dai gestori alle autorità che gestiscono Internet
- Le opzioni gratuite sono una versione limitata di dati commerciali
- Anche questa e' una stima il cui risultato è diverso tra i vari software

Come estraggo i dati.

- AlienIP: http://www.iconempire.com/alien-ip/
- GeoIP: http://www.maxmind.com/app/ip-location
- Geovisite: http://www.geovisite.com/it/
- HostIP.info: http://www.hostip.info/

Superare i limiti presentati

I cookies

- File di testo inviati dal server alla prima connessione.
- Il server può controllare sul browser la presenza o assenza del suo cookie
- Il browser può bloccarli
- L'utente può cancellarli
- Sono abbastanza temuti dagli utenti che li cancellano spesso
- Spesso usati per estrarre dati sulla navigazione in generale da agenzie pubblicitarie.
- Quest'ultima tipologia è molto avversata dagli utenti

Superare i limiti presentati

Uso delle sessioni anonime

- Usabile se invece di pagine statiche si ha un sito con pagine dinamiche
- Tengo nota dell'hand-shake con il gestionale del sito
- Scrivo in un db un identificativo di sessione e ci collego le attività rilevanti
- Chiudo la sessione quando non ho attività da quella fonte che ha fatto l'hand-shake dopo un x temporale (15-60 minuti)
- Preciso nel rilevare gli inizi
- Preciso nel rilevare l'attività
- Leggera imprecisione nella chiusura
- Non lega tra loro le diverse visite
- Non distingue persone/robots

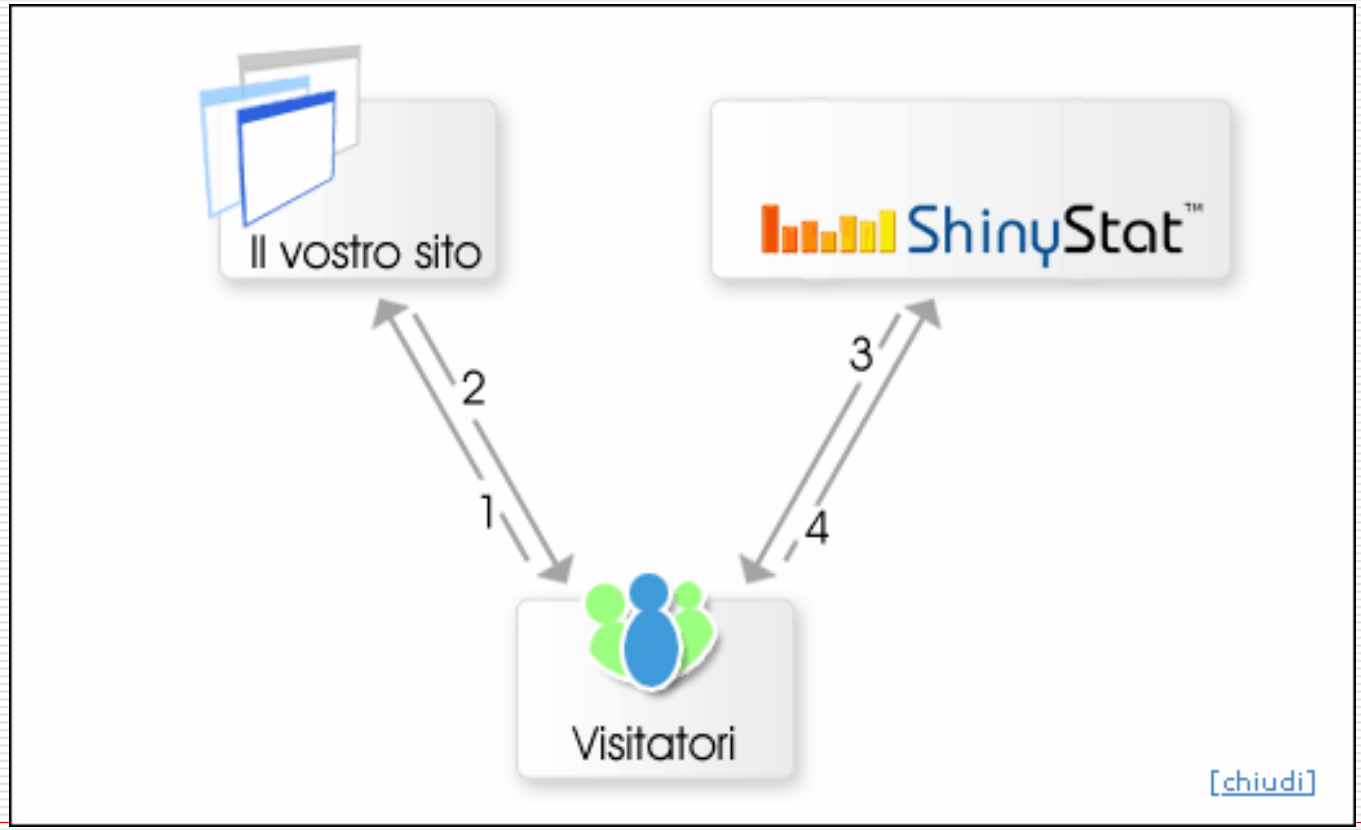
Superare i limiti presentati

Uso sessioni non anonime

- Sempre per siti dinamici
- Sempre con l'appoggio di un db
- Ogni utente e' perfettamente identificato
- Solo persone, no robots
- Le varie visite si collegano tra loro
- Le persone non amano dover ricordare una userid/password
- Voglia di anonimato diffusa
- Non si è presenti sui motori di ricerca [Elsevier ci riesce con accordi ad hoc con google scholar]
- Si esce da quello che è il web normale

Superare i limiti presentati

Portare i dati su una terza parte



Superare i limiti presentati

- Si inserisce del codice javascritpt in tutte le pagine
- E' il browser dell'utente che manda i dati alla terza parte
- I robots non hanno javscripts e dunque sono esclusi
- Può lasciare un traccia sul Pc dell'utente meno invadente dei cookies
- Si usa l'expertise e i mezzi di una grande organizzazione
- Posso esserci problemi di caricamento e di rete usando anche un server di terzi

Superare i limiti presentati

- Google analytics: http://www.google.com/analytics/
- HiStats: http://www.histats.com/
- ShymyStat: http://www.shinystat.com
- Site Meter: http://www.sitemeter.com/

Un esempio

Le statistiche in batch di E-LIS per items

- Punto di partenza: un singolo items
- Es: http://eprints.rclis.org/es/index.php?action=show_detail_eprint&id=6656
- I dati che si vogliono sono:
  - Quanto volte sono stati letti i metadati (views)
  - Quante volte sono stati scaricai i full-texts (downloads)
  - Non considerare i downloads multipli
  - Non considerare i robots

Un esempio

- Si opera sui logs di apache
- Troppo complesso gestire le sessioni anonime
- Si contano le righe con una specifica azione dell'applicazione
- Si fanno de controlli sulla distanza temporale dello stessa azione fatta dallo stesso IP
- Si escludono i robots via lista fissa aggiornata periodicamente

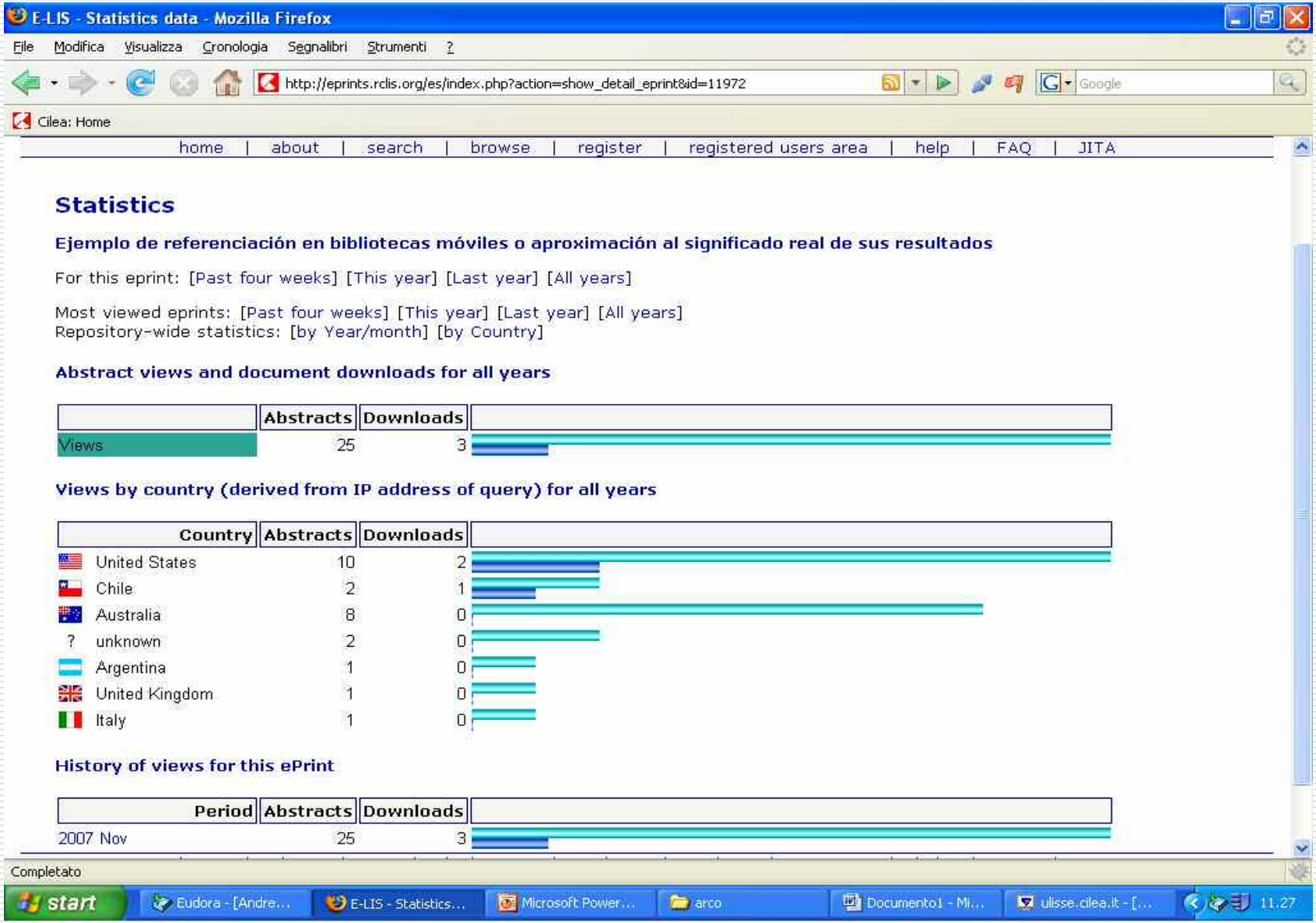
Un esempio

- 213.58.139.xx "-" "-" "- " "GET /archive/00011300/ HTTP/1.1" 200 […]
- Questa è l'indicazione di un view

- 213.58.139.xx "-" "-" "-" "GET /archive/00011300/01/deposito.pdf HTTP/1.1" 200 […]
- Questa è l'indicazione di un download

- Per essere contate due volte le stesse operazioni devono essere distanziate di 180 secondi


E-LIS - Ejemplo de referenciación en bibliotecas móviles o aproximación al significado real de sus resultados - Mozilla Firefox
File Modifica Visualizza Cronologia Segnalibri Strumenti ?
http://eprints.rclis.org/archive/00011972/
Cilea: Home
E-LIS
eprints.rclis.org
E-prints in Library and Information Science
The international open archive for LIS
home | about | search | browse | register | registered users area | help | FAQ | JITA
Ejemplo de referenciación en bibliotecas móviles o aproximación al significado real de sus resultados
Soto Arranz, Roberto (2007) Ejemplo de referenciación en bibliotecas móviles o aproximación al significado real de sus resultados. Boletín de la Asociación Andaluza de Bibliotecarios(83):pp. 11-17.
Full text available as:
PDF - Requires Adobe Acrobat Reader or other PDF viewer.
View statistics for this eprint
Abstract
[Spanish abstract]
Las bibliotecas móviles participan de una serie de factores que, si bien son imprescindibles para su funcionamiento, no pueden considerarse como estrictamente bibliotecarios. Su consideración, cuantificación y análisis son fundamentales para el correcto conocimiento de este tipo de servicios bibliotecarios. Mediante un ejemplo práctico, en este artículo se trata de poner de manifiesto esta circunstancia y reconseguir una aproximación a su verdadero significado.
[English abstract]
Bookmobiles have some factors that are not strictly librarians. But theses factors are very important to know their true meaning. In this text one can see a practical example about it around the comparison of three bookmobiles.
Keywords: Bibliotecas móviles, bibliobuses, referenciación, evaluación. Mobiles libraries, bookmobiles, comparison, evaluation
Subjects: F. Management. > FZ. None of these, but in this section.
Completato
start
Eudora - [Andre...
E-LIS - Ejemplo ...
Microsoft Power...
arco
Documento1 - Mi...
ulisse.cilea.it - [...
11.26


E-LIS - Statistics data - Mozilla Firefox
File Modifica Visualizza Cronologia Segnalibri Strumenti ?
http://eprints.rclis.org/es/index.php?action=show_detail_eprint&id=11972
Cilea: Home
home | about | search | browse | register | registered users area | help | FAQ | JITA
Statistics
Ejemplo de referenciación en bibliotecas móviles o aproximación al significado real de sus resultados
For this eprint: [Past four weeks] [This year] [Last year] [All years]
Most viewed eprints: [Past four weeks] [This year] [Last year] [All years]
Repository-wide statistics: [by Year/month] [by Country]
Abstract views and document downloads for all years
Abstracts Downloads
Views 25 3
Views by country (derived from IP address of query) for all years
Country Abstracts Downloads
United States 10 2
Chile 2 1
Australia 8 0
? unknown 2 0
Argentina 1 0
United Kingdom 1 0
Italy 1 0
History of views for this ePrint
Period Abstracts Downloads
2007 Nov 25 3
Completato
start
Eudora - [Andre...
E-LIS - Statistics...
Microsoft Power...
arco
Documento1 - Mi...
ulisse.cilea.it - [...
11.27

Un esempio

Punti problematici

- La geolocalizzazione
- La lista dei robots non si aggiorna automaticamente
- Notevole spazio necessario sul server
- Soluzione non confrontabile con indicatori di utilizzo sito standard, dunque compresenza di statistiche dovute a webanalizer (a partire da http://eprints.rclis.org/statistics-sheet.html)

DOMANDE ?

Link utili

- "Informatica di base", R. Gaeta, 2004 URL:http://www.di.unito.it/~rossano/DIDATTICA/INF-0304/#Lucidi
- modulo Apache 2 mod_log_config: http://httpd.apache.org/docs/2.0/mod/mod_log_config.html

RFC 2616: http://www.w3.org/Protocols/rfc2616/rfc2616.html

- Sul dynamic DNS: http://it.wikipedia.org/wiki/Dynamic_DNS